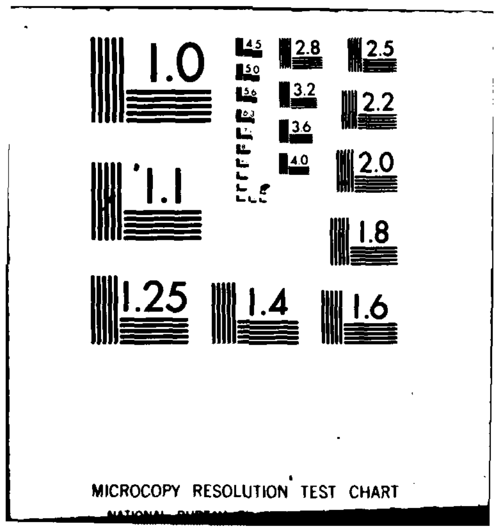
AD-A110 417 NAVAL HEALTH RESEARCH CENTER SAN DIEGO CA F/G 5/2
MONITORING ILLNESS IN A CLOSED WORK ENVIRONMENT.(U)
OCT 81 L HERMANSEN, W M PUGH
UNCLASSIFIED NAVHLTHRSCHC-81-29 NL

END
DATE
FILMED
7-82
DTIC



MICROCOPY RESOLUTION TEST CHART

AD-A110417

MONITORING ILLNESS IN A CLOSED WORK ENVIRONMENT

Larry Hermansen*

and

William M. Pugh*

Naval Health Research Center
P.O. Box 85122
San Diego, California 92138

This article is Report Number 81-29, supported by Naval Medical Research and Development Command, Department of the Navy, under Research Work Unit MF58.524.023-2022. The views presented are those of the authors. No endorsement by the Department of the Navy has been given nor should be inferred.

*Health Care Systems Department

Summary

Illness incidence and illness prevalence are two measures that are often used to assess morbidity rates in a particular population. Illness incidence, a measure of the number of new cases of a particular disorder during a given period of time, is the primary measure used in epidemiological analyses. Illness prevalence is a measure of the number of people ill at any time and therefore is of concern to any manager who needs to keep track of his personnel resources. To compute these indices, it is necessary to distinguish between the initial visit for a medical condition and a followup visit. This distinction becomes obscured, however, when medical data are obtained from encounter forms that are completed during a patient visit, because multiple visits may result from a single condition, or multiple conditions may result in a single visit. In order to reduce such ambiguity, a model of illness etiology was articulated and developed into a computer program. This program converted data on individual patient visits to meaningful records of illness or injury episodes. Illness incidence was then computed by counting these episodes. To compute illness prevalence from the model, it was necessary to incorporate estimates of illness duration for specific diagnoses. Incidence could then be combined with duration to determine how many people were ill at any given time. Finally, each measure—illness incidence and illness prevalence—was plotted so that the two measures could be compared and temporal trends observed.

A prime requisite in the study of disease and the management of personnel in a work environment is the ability to assess accurately the morbidity in a given population. To meet this need a system designed to provide an accurate and efficient method of monitoring illness was developed and tested aboard deployed U.S. Navy ships for a six-month trial period (1, 2). An integral part of this system was a medical treatment reporting form which provided basic demographic and medical treatment information. The form was designed so information provided could be electronically tallied via an optical scanning instrument. This system represented a vast improvement over former methods of monitoring outpatient illness rates in a closed work environment. This paper presents additional procedures which were used to further organize and interpret these data to make them more meaningful to the researcher, health care practitioner, and production manager.

While tallies of data in the experimental medical data reporting form provide a measure of illness visits, this is only one index of morbidity. As MacMahon, Pugh, and Ipsen (3) point out, the investigator studying the origins and spread of a particular disease is primarily concerned with the number of new cases in a population during a given time period, or illness incidence, and would not be interested in return or follow-up visits. A production manager, however, is likely to be more concerned with the number of people indisposed due to illness or injury at a given time, or illness prevalence, because that figure provides a better indication of reductions in the work force due to illness. However, because illness incidence and illness prevalence are indices reflecting the health status of a population at specified points in time they cannot be directly measured by individual medical treatment forms which are filled out only when a person seeks treatment through a health care professional. Instead it is necessary to further process the individual visit data using means that will be described in this paper.

Methods for constructing illness incidence and prevalence indices from data provided on medical treatment reporting forms will be described. The paper will also demonstrate how data aggregation, temporal sequencing, and disease modeling can be used to derive the population indices. In addition, processes that otherwise often remain covert in the creation of such indices will be demonstrated.

Illness Incidence

Problems of Measurement

On the surface it would appear that simply requiring, on each treatment form, an indication that a visit was either an initial visit or a follow-up would provide sufficient information to compute the illness incidence measure since all follow-up visits could be ignored. Problems, however, have been shown to arise when using this method. Requesting either the patient or the health care practitioner (hospital corpsman) to indicate whether a visit was "initial" or "follow-up" is helpful, but sources of inaccuracy still exist. For example, a patient or corpsman relying on memory may forget a previous visit, so that two initial visits may be recorded for the same disorder. On the other hand, a patient with two visits for different complaints may attribute them to different disorders while the corpsman may recognize both as symptoms of a single disorder. Finally, multiple independent disorders may be treated and recorded during a single visit. Therefore, when a visit is recorded as initial visit but multiple complaints are indicated, there may be ambiguity about which of the different disorders are new.

To reduce the effects of such factors when computing illness prevalence a rudimentary model of illness etiology was applied. The model used is generally consistent with concepts

of disease developed by Fabrega (4) who stated that "disease is a temporally extended undesirable deviation of a human characteristic or set of characteristics (p. 125)." In addition, Fabrega indicated that for each disease there is a set of defining characteristics and that there also is a set of indicators which allow one to identify or classify a disorder. It is necessary, however, to develop the above concepts further to obtain sufficient detail for their application to a specific task.

Definitions

The first step in developing the model is to define the term used to label various components. Thus, in the following discussion the period from the initial appearance of illness (or decrement in health status) to recovery is termed an illness episode. Other key terms used to explicate this model are: symptom, diagnosis, cluster, and temporal contiguity. A symptom is defined as a physical manifestation or a subjective sensation or perception that results from some underlying disorder. A diagnosis is the identification of a disease state or malfunction and the causal agents which precipitated it. A cluster is composed of a set or series of symptoms and/or diagnoses that are causally related. For example, two disorders would be within a single cluster if they have common causes, or if one causes the other. Finally, causal association may be inferred from temporal contiguity; that is, symptoms or diagnoses that are in close proximity timewise and of the same cluster as to be considered causally related.

Identifying illness episodes

Using the above terms, a model was formulated to identify illness episodes based upon data obtained from medical treatment reporting forms. Once separate episodes are identified, illness incidence can be measured by counting the number of illness episodes. As previously noted, the distinguishing feature of an illness episode is an undesirable change in health status that is recorded. If an individual has any visits recorded, it is assumed that at least one illness episode has occurred. But when a patient has made multiple visits during some time interval, it must be determined whether each succeeding visit was the continuation of a prior condi.ion or the onset of a new illness. Two visits will be considered to be the result of a single illness episode if they are temporally "close" and if they are within the same cluster. For example, a patient who is treated for an upper respiratory infection at one visit and for pneumonia two days later would have both visits assigned to a single illness episode. However, if two visits are distant temporally or are not from the same cluster (e.g., pneumonia and fracture), they would be assigned to different illness episodes.

After illness episodes are identified, they are labeled with a symptom or diagnostic code. The procedures for selecting the episode label are designed to result in the most descriptive and appropriate code being used. Of course, if there is only one visit in an episode, the episode is labeled with the same code used to classify the presenting complaint. In the case of an episode with multiple visits, the label applied depends upon additional circumstances. In those cases where all visits are classified with the same code, or where only one of the visits received a diagnostic code, there is little ambiguity. The situation is more complicated when more than one diagnosis from one cluster is used, however. If one of the conditions cited is more specific and descriptive than the others, then it is used to label the episode. Otherwise, the last complaint that occurs in the series is used to label the episode.

Application

The system designed to generate illness episode information from dispensary visit records involved the development of a computer program which incorporated the above logic. To implement this program, a series of judgments had to be made explicit. First, each code used to label a treated complaint was designated by the user as being "diagnostic" or "symptomatic." For example, one probably would consider appendicitis as a diagnostic label and complaints such as headache or fever as symptoms. Second, interrelated disorders (symptoms and/or diagnoses) had to be grouped to form illness clusters. Third, the time interval considered to be "close" and "distant" had to be quantified so that the notion of temporal contiguity could be implemented. That is, visits occurring close together will be treated as part of the same episode while visits more distant in time will be counted as separate episodes. If multiple visits are neither "close" nor "distant," then other information on the medical treatment reporting form, that is, "initial" versus "follow-up", must be used to decide whether the second visit should be included in the same episode as the first. Fourth, the user is required to assign a priority level to the various diagnostic categories. This indicator is used to decide how to label an episode which includes two or more visits for complaints classified as "diagnostic." Thus, labels for disorders assigned high priority values are treated as though they were more specific or descriptive than the terms assigned low priority scores.

The system described thus far for converting raw data from the medical treatment forms into illness episode information is diagrammed in Figure 1. Central to this system is a FORTRAN program which incorporates the logic representing the model of illness etiology. The input for this program includes not only the medical treatment form data but also a set of parameter cards which allow the user to exercise his judgment regarding what is a symptom or diagnosis, what is close or distant, and so on. As indicated in Figure 1, the combination of the model of illness etiology and the user's judgments results in the operational definition of an illness episode. By applying this definition to the raw data for illness visits, a file of illness episodes is created, and a simple tally of these illness episodes yields an estimate of illness incidence.

In addition to fulfilling the primary objective of converting illness visit information into illness episode data, the above system is beneficial in other ways. For instance, coding errors and logical inconsistencies may be detected and corrected while the data are being processed. Another benefit of this system is that it makes the process of converting the raw data into records of illness episodes explicit so that covert and possibly inconsistent procedures are avoided.

Illness Prevalence

Computation

As noted previously, illness prevalence or the number of individuals ill or indisposed at any given time is probably the most relevant morbidity indicator from a management perspective. Illness prevalence cannot be computed directly from the medical treatment form data, however. MacMahon et al., (3) note that illness prevalence is a combination of illness incidence and illness duration (length of an episode). Thus, the model used above must be extended to include illness duration before prevalence can be estimated.

One way to accomplish this is to assume that a person seeking medical attention on a particular day felt ill prior to the visit and will continue to be ill for some time after treatment. Then prevalence could be determined by counting the number of new illness episodes occurring each day and assigning an equal number to the surrounding days. For example, if 10 people sought medical attention for respiratory ilness on one day and the course of the disease

(time from onset to full recovery) is typically seven days, then in addition to the day of the visit, those 10 people would be counted as being ill the 3 days prior to and the 3 days following their visit.

Use of Moving Average

The above method for computing illness prevalence may be implemented using a method developed for smoothing temporal trends in time-series analysis (5, 6) although the rationale behind its use is somewhat different. In time-series analysis temporal sequences are often smoothed by a moving average process where the score for one period is weighted. Each weighted score is then added to adjacent values and the total for each period is divided by the sum of the weights. The rationale behind these procedures is based on the notion that the data at each point in time contain some error but that data from contiguous times can be used to help estimate the true value for the period in question. Thus, by using the actual data for a given point in time as one estimate of the true value and combining that with estimates from surrounding data points, error can be reduced, thereby exposing the underlying trend. However, the objective in this paper is not to combine and average estimates of some true value. Instead, the surrounding data are viewed as a legitimate part of the circumstances existing at one point in time. That is, when a person seeks treatment for influenza on one day, it is likely that he will not be fully recovered on the next day and therefore he will be counted as being ill for both days. Thus, a moving sum procedure was used to compute prevalence which is the same as the moving average procedure except that the weighted sums of temporal sequences were not divided by the sum of the weights. Therefore, in the earlier example in which 10 people had respiratory illnesses that lasted seven days each, the procdures described were those used to compute a simple moving average except that the result was not divided by the sum of the weights.

Results

Data Edits

A computer program was designed to implement the logic developed in this paper and was applied to medical treatment report data for the crew of an amphibious assault ship. Data for each visit were edited to be consistent with information about other visits. Although the majority of records remained unchanged, Table 1 shows a few examples of patients' illness data prior to editing as well as the changes that were made. Most changes were of the type shown for Cases 1 and 2 where follow-up visits apparently had been recorded incorrectly as initial visits. The next most frequent type of change is exhibited by Case 3 in which visits that had been recorded as follow-up visits had no preceding initial visits within a reasonable time frame.

Cases 4 and 5 show a combination of changes within a single record. Case 5 is a particularly interesting record. First, it provides an example of a symptom of "hazy diagnosis," (unspecified Respiratory Diseases) preceding a more clear-cut diagnosis, Upper Respiratory Infection (URI). Case 5 is also interesting because of the pattern of illness visits occurring between January 26th and February 4th. Each visit was recorded as an initial visit by the corpsman; however, considering the type of complaints and their contiguity, one might suspect that the patient actually had a single influenza episode.

Data Smoothing

After the illness record for each patient was edited, daily incidence of illness was computed by finding the number of initial visits for each day and prevalence was computed as described earlier. Then to demonstrate the effect produced by each procedure, daily incidence

and prevalence of respiratory illnesses incurred during the deployment were plotted with the expected duration of a respiratory illness episode fixed at seven days. The graph generated by plotting the incidence data is shown in Figure 2 and Figure 3 shows how these data appear after using the illness prevalence transformation..

In these Figures the Y axis shows the percentage of the crew that was affected by respiratory illness. Each character along the X axis represents one day. Alternate strings of "As" and "Bs" are used to indicate the months of the year with the initial string of "Bs" representing the latter half of November, followed by a string of As for December, a string of Bs for January and another string of As for the first part of February. The values for each day are printed as a column of "Ps" or "Ss" where P indicates that a ship was in port on a particular day and S indicates that the ship was at sea.

Comparing the illness prevalence estimates shown in Figure 3 with the incidence data shown in Figure 2, it becomes clear that, at any one time, illness prevalence is much greater than the proportion who seek medical attention. Therefore, it is felt that for anyone concerned with the effect that the illness within a certain population may have on production or mission effectiveness, this type of data transformation and display could be quite useful.

Discussion

Even though the procedures used in the present paper are in the preliminary stages of development, they appear to greatly enhance analysis of illness data. Modifications to individual records in most cases were straightforward. For example, it is not difficult to justify counting three visits within one week for pharyngitis as a single illness episode rather than three. Some individuals, however, had complex illness patterns which suggested an underlying diagnosis that may have eluded the corpsman. In the future, it may be possible to use more sophisticated procedures to identify meaningful illness clusters and syndromes.

Whenever data are modified, there is a question about the validity of the changes and this is a question that could and should be investigated in future studies. With respect to illness incidence, data were modified only when two or more points were clearly inconsistent and then a best guess type strategy was employed to resolve the discrepancy and render the data meaningful. Thus, it is believed that the overall amount of error was reduced but the lack of a second source of illness data prevents one from obtaining a conclusive answer to this question. Perhaps, in future systems these methods could be used to alert health care personnel of discrepancies in the data obtained so that immediate steps could be taken to resolve the problem.

Finally, it is believed that the techniques described here along with the traditional time-series approach, can be used to form a more comprehensive picture of illness and injury patterns particularly in an industrial environment where trend fluctuations, seasonal variations, and irregular effects are so important to the production manager for future planning.

References

1. Hermansen, L.A., Jones, A.P., & Butler, M.C. Development of an outpatient medical treatment reporting system for shipboard use. U.S. Navy Medicine, 71, 16-21 (1980).
2. La Rocco, J.M., & Gunderson, E.K.E. A proposed new outpatient data collection system. Report No. 78-9, Naval Health Research Center, San Diego, California (1978).
3. MacMahon, B., Pugh, T.F., & Ipsen, J. Epidemiological Methods. Boston: Little, Brown, & Co. (1960).
4. Fabrega, H. Disease and Social Behavior: An Interdisciplinary Perspective. Cambridge, Mass.: The MIT Press (1974).
5. Anderson, T.W. The Statistical Analysis of Time Series. New York: John Wiley & Sons (1971).
6. Kendall, M.G. Time-series. London: Griffin & Co., Ltd. (1973).

Table 1

Typical Modifications to Individual Illness Records

Original Data

Case	Date	Visit	Complaint	Modification
1	27 Jan	Initial	Pharyngitis	-
	28 Jan	(Initial)[a]	Pharyngitis	Follow-up
	1 Feb	(Initial)	Pharyngitis	Follow-up
2	12 Nov	Initial	Gonorrhea	-
	15 Nov	(Initial)	Gonorrhea	Follow-up
	17 Nov	Follow-up	Gonorrhea	-
	22 Feb	Initial	URI	-
	6 Mar	Initial	Gastritis	-
3	19 Oct	(Follow-up)	Pyorrhea	Initial
	5 Nov	Initial	Gonorrhea	-
	20 Dec	(Follow-up)	Musculoskeletal	Initial
	22 Dec	Initial	Diarrhea	-
	5 Jan	Follow-up	Musculoskeletal	-
	10 Jan	Follow-up	Musculoskeletal	-
	17 Feb	(Follow-up)	Musculoskeletal	Initial
4	12 Oct	Initial	Otitis Externa	-
	30 Nov	(Follow-up)	Open Wound	Initial
	14 Jan	(Follow-up)	Open Wound	Initial
	24 Feb	Initial	URI	-
	3 Mar	(Initial)	URI	Follow-up
5	26 Jan	Initial	URI	-
	26 Jan	Initial	Motion Sickness	-
	27 Jan	(Initial)	URI	Follow-up
	4 Feb	Initial	Diarrhea	-
	5 Mar	Initial	Skin Disorder	-
	5 Apr	Initial	(Unspecified Resp. Disease)	URI
	7 Apr	(Initial)	URI	Follow-up

[a]Parentheses indicate the original data that were subsequently modified.

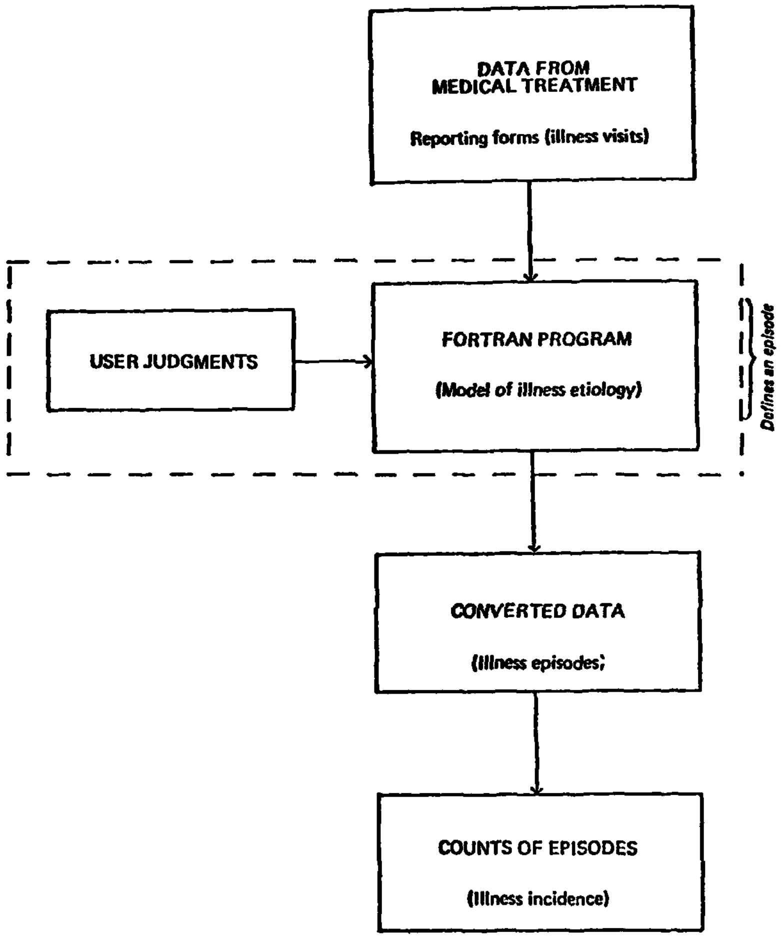


Fig. 1. System for Computing Illness Incidence from Records of Illness Visits.

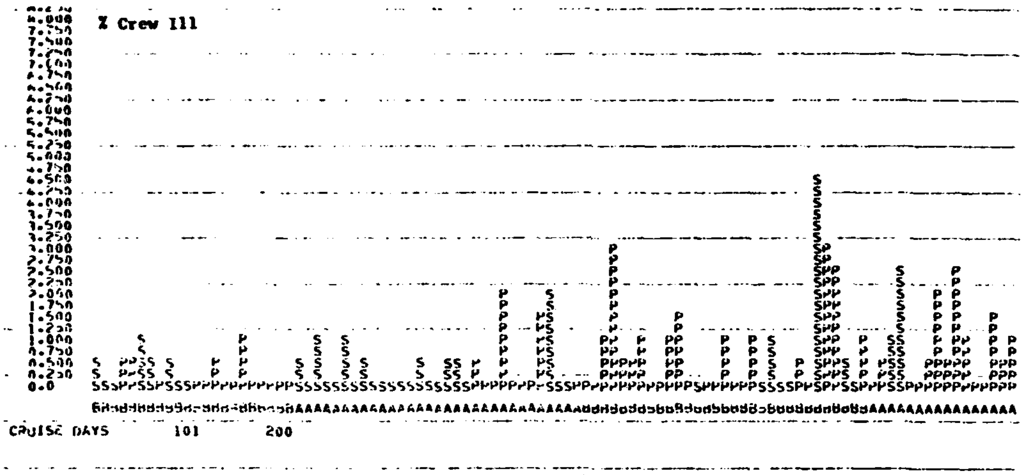


Fig. 2. Daily Incidence of Respiratory Illnesses

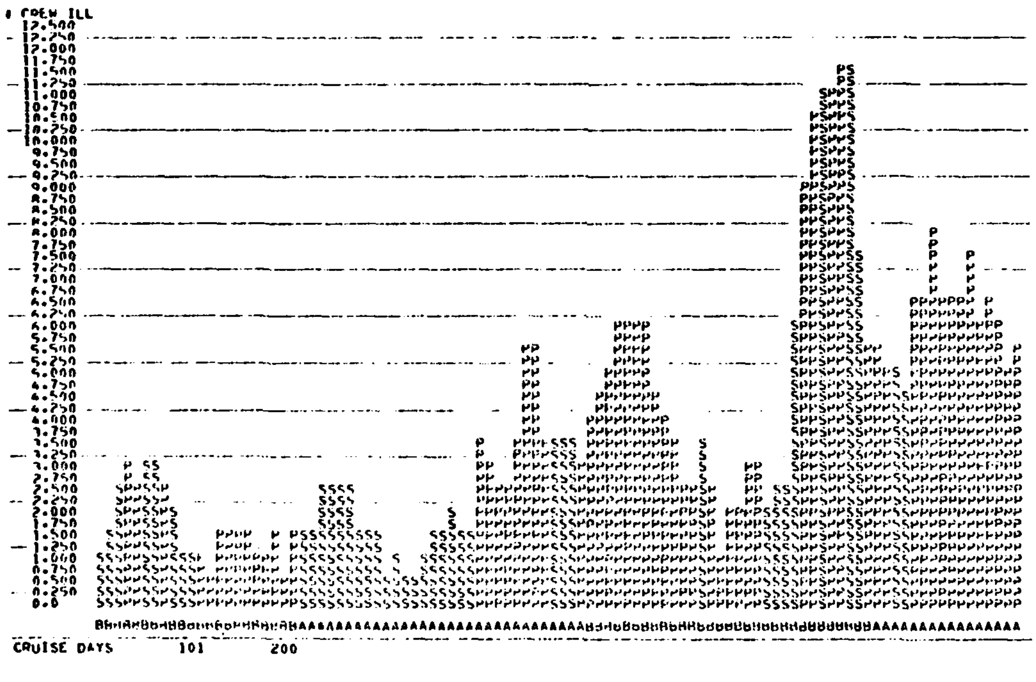


Fig. 3. Prevalence of Respiratory Illnesses

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 81-29	2. GOVT ACCESSION NO.: AD-A110 418	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): Use of an Automated Patient Record Aboard Ship		5. TYPE OF REPORT & PERIOD COVERED: Interim
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Larry Hermansen and William M. Pugh		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Naval Health Research Center, P.O. Box 85122, San Diego, CA 92138-9174		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: MF58.524.023-2022
11. CONTROLLING OFFICE NAME AND ADDRESS: Naval Medical Research & Development Command, Bethesda, MD 20014		12. REPORT DATE: 20 Oct 1981
		13. NUMBER OF PAGES: 10
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): Bureau of Medicine & Surgery, Department of the Navy, Washington, DC 20372		15. SECURITY CLASS. (of this report): UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

Presentation at the Fifteenth Annual Hawaii International Conference on System Sciences

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

medical records, illness incidence, illness prevalence, illness episode

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In the ship's dispensary medical data are typically gathered on individual patient visits. Because multiple visits may result from a single condition or multiple conditions may result in a single visit, some overt or covert process must be used to develop records of discrete illness episodes with accurate diagnostic information. In an effort to make this process explicit, a model of illness etiology was articulated, and a program based upon that model was developed. This program converted data on individual patient visits

into meaningful records of illness or injury episodes. In addition, estimates of illness duration for specific diagnoses were incorporated into the model thereby allowing illness prevalence to be computed and plotted in terms of the percentage of the crew ill in any time period. Such a system should aid in the management of personnel resources.